ED101280

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RETRIEVAL OF SENTENCE RELATIONS: SEMANTIC VS. SYNTACTIC DEEP STRUCTURE

Charles A. Perfetti

Learning Research and Development Center
University of Pittsburgh

1973

To appear in the International Journal of Cognition.

The research reported herein was supported by the Learning Research and Development Center, supported in part by funds from the National Institute of Education (NIE), United States Department of Health, Education, and Welfare. The opinions expressed do not necessarily reflect the position or policy of NIE and no official endorsement should be inferred.

CS 001 564

DOCUMENT RESUME

ED 101 280 CS 001 566

AUTHOR Perfetti, Charles A.
TITLE Retrieval of Sentence Relations: Semantic vs. Syntactic Deep Structure.
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO LRDC-1973/9
PUB DATE 73
NOTE 22p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Cues; Deep Structure; Higher Education; *Language Research; *Paired Associate Learning; *Semantics; *Sentence Structure; Surface Structure; *Syntax

ABSTRACT

Two experiments on unaided and cued recall of sentences presented in context to college students are reported in this study. Key nouns in the sentences were arranged to have uniform surface functions, but to vary independently in deep syntactic category and semantic function. Cued recall for sentences in which the semantic function of actor and recipient coincided with the syntactic function of deep subject and object, respectively, was better than for sentences which did not have this normal semantic-syntactic coincidence. Unaided recall was not different for the two types of sentences. Models of sentence processing may have to represent both types of information as available to the language user. (Author)

Abstract

Two experiments on unaided and cued recall of sentences presented in context are reported. Key nouns in the sentences were arranged to have uniform surface functions, but to vary independently in deep syntactic category and semantic function. Cued recall for sentences in which the semantic function of actor and recipient coincided with the syntactic function of deep subject and object, respectively, was better than for sentences which did not have this normal semantic-syntactic coincidence. Unaided recall was not different for the two types of sentences. Models of sentence processing may have to represent both types of information as available to the language user.

RETRIEVAL OF SENTENCE RELATIONS: SEMANTIC VS. SYNTACTIC DEEP STRUCTURE[1]

Charles A. Perfetti
University of Pittsburgh

One of the psychologically significant aspects of transformational grammar is the representation of relational information that is not directly revealed in sentence surface structure. Studies by Blumenthal (1967) and Blumenthal and Boakes (1967) were especially important on this point, because they showed clearly a divergence between deep and surface structure in either the storage or retrieval of sentences. The main result of these studies was the demonstration that in sentences which were superficially similar, but different in deep syntax, differences in the probability of words from the sentences to cue recall of the entire sentences were associated with their deep structure roles. For example, comparing the two sentences The officers were eager to please and The officers were easy to please, officers proved to be a more potent prompt in the first sentence, where it is deep subject, than in the second sentence, where it is not.

There are, however, semantic relations that correspond in many cases to the syntactic relations that are revealed in transformational grammar. For example, most noun phrases which are subjects in deep structure are also actors or agents semantically. Thus, in Blumenthal's sentences, and in most examples from the literature of transformational

---

[1]The research reported in this paper was carried out with the substantial assistance of Blaine Garson, who collected the data, and Robert Lindsey, who assisted with judging responses and analyzing data.

grammar, the deep subject also names the actor of the action described in the sentence. This is the case in the above sentences, for example.

However, this strict correspondence between the semantic notion of actor and the syntactic notion of deep subject does not always hold. Of course, there are many sentences that do not have action as the semantic property. Statives such as Horace is sympathetic are only one of many semantic relations other than that of action. More to the point, there are sentences which have an action semantic, but for which the actor cannot be identified with deep subject. For example, in a sentence such as Virgil experienced humiliation at the hands of the soldiers, it is not Virgil but the soldiers who comprise the actor; Virgil, however, is the subject.

It was sentences of the latter type which were of interest in the present study. The immediate empirical question is whether, in sentences such as the one above, Virgil or the soldiers is the more potent cue for retrieving the entire sentence (when compared to a sentence in which Virgil is both actor and the deep subject). The more general question is whether semantically based relationships such as action (actor, recipient) are cognitively more significant than syntactic relations such as subject of. Case grammars (Anderson, 1971; Fillmore, 1968) assign theoretical importance to these semantic concepts (although as parts of syntactic theory), and thus provide interesting alternatives to phrase-structure based grammars.

Method

Sentences

The experiments were designed to test the cued recall of sentences under four conditions which varied according to specific syntactic and semantic properties of the noun which served as a cue. The noun cue was: (1) Deep Subject and Actor (SA), (2) Deep Subject and Recipient of Action (SR), (3) Deep Object and Actor (OA), and (4) Deep Object and Recipient of Action (OR).[2] These conditions are shown in the following sentences:

(a) The MAYOR publicly denounced the policies of the GOVERNOR.

(b) The MAYOR withstood great pressure from the GOVERNOR.

(c) The DEFENDANT reluctantly told his story to the PROSECUTOR.

(d) The DEFENDANT experienced embarrassment from the questions of the PROSECUTOR.

Sentences (a) and (b) form a paired experimental observation, and (c) and (d) form another. Sentences such as (a) and (c), in which actor and deep subject coincided, were called "Normal." Sentences in which the actor was a syntactic object, such as (b) and (d), were called "Marked." The capitalized nouns are the cues for recall and are the same words for both

[2] The "deep object" varied somewhat in its surface relations, occurring, for example, sometimes as the object of a preposition and sometimes as the indirect object of a verb. Hence, its exact role in deep structure was not uniform, and it was sometimes the subject of a sentence embedded in the verb phrase. Its defining attribute was that it occurred in deep structure in a constituent dominated by VP.

members of each pair. There were 10 pairs of 20 sentences constructed for the experiments.

The cueing conditions are illustrated as follows: MAYOR is SA in (a) and SR in (b); GOVERNOR is OR in (a) and OA in (b). Two lists were constructed, so that, for each sentence pair, the SA sentence occurred in one list and the OA occurred in the other. Half of the sentences in each list were cued by NP1, and the other half by NP2. By comparing the effectiveness of a cue word under its two conditions, the relative contribution of the two controlled cueing properties is obtained.

Other noteworthy characteristics of the sentences include the following:

1. The surface subject position was not varied. The first noun phrase (NP1) in surface structure was always deep subject. NP2 had various surface grammatical functions, but was never deep subject.
2. There was a slight difference in the length of the two types of sentences comprising the experimental pairs. Normal (SA) sentences, exemplified by (a) and (c), averaged 8.9 words per sentence (range: 7-10), while Marked (OA) sentences, exemplified by (b) and (d), averaged 9.9 words (range: 7-12). The slight length difference was mainly due to an additional grammatical word required by sentences in which the deep subject is not the actor, and thus the lexical density (Perfetti, 1969) of the types was about equal, .54 for SA and .52 for OA. Variations within

types are of little consequence, since the main comparison was to be between the relative cue potency of the two words from the same sentence compared across the two sentence types.

Experiment I: Subjects and Procedure

Twenty-four University of Pittsburgh undergraduates participated in Experiment I, twelve for each of the two lists. Each S was tested individually.

Since the purpose was to test the retrieval of meaningful relations among sentences, as well as verbatim recall, it was important to ensure that the meaning of the sentences was processed. To this end, a procedure similar to that used by Blumenthal (1967) was used. Each experimental sentence was part of a "brief excerpt from a story" which contained three sentences, the last one being the test sentence. E read the first two sentences to S, and then showed S a card with the third sentence which was then read aloud by S. Instructions to S were that he would be tested for his understanding of the story excerpt and for his recall of the final sentence. The following is an example of an excerpt heard by S: The governor was the most powerful political figure in the state. The mayor was known to be independent. The mayor publicly denounced the governor or The mayor withstood great pressure from the governor. All Ss heard the first two, while the final sentence varied according to whether it was an SA or OA condition. The above example illustrates a significant characteristic of the excerpts: Each noun that was to be a cue occurred exactly once in the preceding context sentences and always in the subject position. (The Appendix lists all the test and context sentences.)

The input phase of the experiment was followed by a five-minute interval in which S made judgments of line drawings. Then followed three tests:

1. Free recall. S was asked to write in a booklet of blank pages each key sentence that he had heard, one sentence per page.
2. First cued recall. S was now required to go through another booklet, this time writing a sentence next to its cue word, one per page. For example, the word Governor would be in the booklet for one S, while another S would respond to Mayor.
3. Second cued recall. Finally, S went through a third booklet which had the alternate cue from each sentence. If Governor had occurred in the first cued recall test, Mayor occurred in the second.

Experiment II

In order to observe a direct relationship between the two nouns of the sentence, a second experiment was conducted. The materials were identical to those of Experiment I, and the procedure differed only in that Ss were asked to recall only the other noun from the sentence given one noun as a prompt. This is essentially a replication of the two prompted recall conditions of Experiment I for the case in which S does not have to produce the entire sentence. Recall data should be a fairly direct indication of the stored connection between the two nouns. Twenty-four Ss were assigned to cueing and sentence type conditions in the same fashion as in Experiment I.

Results and Discussion

Recall Measures

In these experiments, the concern was with the retention of meaning rather than with exact reproduction of the sentence. Accordingly, responses were considered to deviate from perfect reproduction in an ordinal manner which could be reliably scored and which roughly corresponded to an underlying scale of meaning retention. Thus, a response was placed into one of the following categories, ordered from least to most retention:

1. Absent or no response other than the prompt. (A)
2. Partial recall. Meaning not preserved. (P)
3. Inference recall. Meaning not preserved, but recall implicationally related to target sentence. (I)
4. Meaning-preserving recall. (M)
5. Verbatim recall. (V)

Categories (M) and (V) were of primary interest since they are measures of meaning-preserving recall. The criterion for (M) was that the recall directly revealed the relationship between the actor and the recipient; that is, the relationship ACTION (ACTOR, RECIPIENT) was maintained without respect to perfect lexical or syntactic recall. The categories are cumulative in the sense that any higher category included the amount of meaning preservation reflected in a lower category. Specifically, meaning-preservation recall (M) included both categories (M) and (V). (I) and (M) do require judgment in scoring, and they are illustrated in the following response protocol:

Target sentence: The artist suffered insults from the pen of the critic.

(P) The artist drew the pictures.

(I) The critic gave a bad review to the performance.

(M) The artist was insulted by the pen of the critic.

It is easy to see that (P) is well off target, but the (I) and (M) judgments do require some explanation. In (M), although the words are not quite all there, it is clear that the relationship among elements in the target sentence has been preserved, viz. INSULT (ARTIST, CRITIC). In the case of (I), this is not quite the case. What has been produced is a relationship that is consistent with the meaning of the target sentence; if the artist was insulted at the pen of the critic, then it is quite consistent psychologically to infer that the critic gave some "performance" a bad review. It does not quite preserve meaning because it fails to reflect INSULT (ARTIST, CRITIC).

In general, word substitutions which did not change meaning relationships in the sense outlined above were counted as (M). Sentences which failed to produce the relationship but were otherwise inferentially consistent with the target were (I) and those that were neither were (P). Despite the potential difficulty of the classification task, there were few problems and very good agreement in classifying. The interjudge agreement between two judges was 90 percent overall and 95 percent when based only on whether a given sentence was to be (M) or better vs. (I) or worse. In all subsequent reports of results, the measure is the cumulative (M) category of (M) plus (V), except where noted.

Experiment I

The measures described above were applied to all response attempts of the first experiment. Cumulative percentages for three of the categories are shown in Table 1. These data show that as the criterion

is relaxed, the relationship between sentence type and retention remains essentially unchanged. The nature of this relationship is most clearly seen in Figure 1, which shows the percentage of (M) responses for successive recall attempts for both Normal and Marked sentences.

Table 1

Recall Data: Experiment I

Sentence Type	Condition	Cumulative Recall Categories		
		(V)	(M)	(I)
Normal	Unaided	.05	.23	.32
	Prompt N1	.10	.55	.63
	Prompt N2	.17	.68	.72
Marked	Unaided	.03	.23	.30
	Prompt N1	.05	.42	.53
	Prompt N2	.07	.43	.58

Note. The column categories represent successively relaxed response criteria with cumulative proportions. The prompting conditions are for the first prompting only.

There are two rather interesting aspects to Figure 1. For one, Normal sentences are clearly better retained than Marked sentences; but, for the other, this superiority is only under conditions of prompting. When Ss are initially asked to recall all sentences, they are equally able to produce the meaning of Normal and Marked sentences. However, when first one noun and then the other is used to prompt the subject, there is a significantly greater improvement for Normal sentences compared to Marked sentences. The low-level conclusion is that somehow both subject and

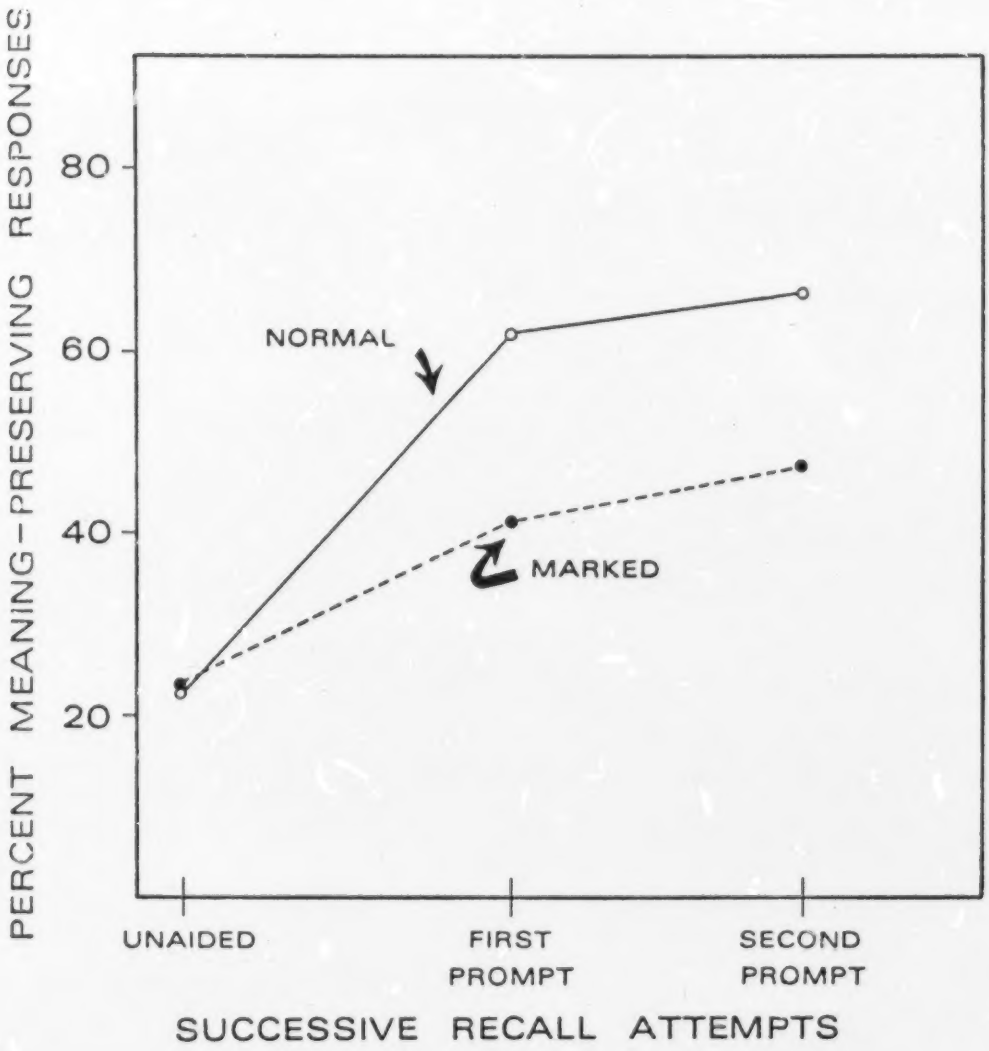

Figure 1. Retrieval of Meaning on Successive Recall Attempts.

object (or actor and recipient) nouns provide better access to the meaning of sentences in which the actor and the subject are the same than to sentences in which the actor is not the subject.

Data on the other variable indicate that it did not much matter which noun was used as a prompt. Here the measures are taken only on the first prompted recall trial, since the interpretation of a second prompted recall is rather difficult. The improvement in (M) recall from the unaided trial to the first prompted trial actually provides the best indication of cue effectiveness. On the first prompted recall, for Normal sentences, the proportion of (M) responses was .55 when the first noun was the prompt and .68 when the second noun was the prompt. For Marked sentences, these figures were .42 and .43, respectively. A two-factor analysis of variance for repeated measures showed that this (M) improvement measure was affected significantly by sentence type (Normal better than Marked; $F = 5.27$, $p < .05$), but only marginally by prompt (first noun vs. second noun; $F = 3.20$, $p < .10$). The interaction was not significant ($F < 1$). A similar picture emerges if, instead of improvement, absolute performance on the first prompting trial is measured, except that the error variance is somewhat greater, and the prompt factor has an F of less than one.

Experiment II

Experiment II was designed to get a relatively simple measure of noun recall given noun prompt, without requiring S to produce the whole sentence. Thus, on the one hand, it provides a measure of what characteristic of the noun (syntactic category or semantic function) is important for its prompting effectiveness; on the other hand, it provides a type of control for the first experiment. That is, are differences due to sentence

type and prompt strictly related to the production of meaning in recalling sentences, or are they present in noun-noun recall? It is possible, of course, to get similar data for Experiment I. First in unaided recall, what is the probability that one noun is recalled given that the other noun was also recalled? In prompted recall, the measure is the presence of the other noun, or an acceptable substitute, given one noun as a prompt while ignoring the remaining recall content. This is strictly analogous to the data of Experiment II.

The conditional probabilities are shown in Tables 2 and 3. Table 2 shows the unaided recall trial of Experiment I where the measure is the presence of the exact noun or a meaning-preserving substitute. (This liberal measure is required by the fact that the main interest in this research is in meaning preservation not verbatim recall, although the relative figures are the same when the verbatim requirement is imposed.) The first two columns show that there was virtually no difference in the probability of recalling the first versus the second noun, and no difference for Normal versus Marked. The last two columns show the very high degree of noun integration present in both types of sentences. Recall is all or none in the sense that if one noun is recalled, both are recalled. Differences between (N2/N1) and (N1/N2) are not significant. It is important to note this all or none characteristic, since the integration of a memory unit, as indicated by conditional recall probabilities of elements within the unit, has been related to what part of the unit is a good prompt for the whole unit (Horowitz & Prytulak, 1969). Specifically, for a well-integrated unit, the best prompt is the most available part of the unit, whereas for a poorly integrated unit, the best prompt is the least available part. Integrated recall is also evidence for this experiment when the whole sentence is considered. In the unaided condition, the recall of a whole sentence which preserves meaning or is implicationally related

Table 2

Unaided Noun Recall: Experiment I

Sentence Type	Noun Type (N1)	(N2)	(N2/N1)	(N1/N2)
Normal	.38	.36	.85	.93
Marked	.41	.38	.91	.98

Note: The cell entries are proportions. The conditionalized proportions in the right half of the table are based on the recall of the first-named noun given that S recalled the second-named noun.

Table 3

Prompted Noun Recall

	(N2/N1)	(N1/N2)	Average
Experiment I			
Normal	.70	.77	(.735)
Marked	.70	.73	(.715)
Experiment II			
Normal	.72	.78	(.750)
Marked	.72	.75	(.735)

Note: The first column is recall of N2 given N1 as a prompt. The second column is the inverse.

to the target is about 74 percent given any recall at all. And this underestimates the degree of integration in the sense that it excludes many whole sentences which are sensible and related to the original context of the target sentence.

In Table 3, the conditional prompted recall of nouns alone for both Experiment I and II can be seen. The fact that in Experiment I whole sentence recall was required made no difference. When prompted to recall the whole sentence, S was as able to produce the other noun from the sentence as when he was asked to produce only the other noun. Since instructions at input were the same in the two experiments, it is a compelling conclusion that the stored meaning of the sentence mediated noun recall in Experiment II. Subjects recalled the sentence given the noun prompt, and then produced the other noun.

There is no significant difference owing to sentence type, and there is no significant difference owing to prompt type ($F < 1$). The pattern for prompt type is the same as for the (M) measure but with the difference even smaller.

Conclusion

The results of these two experiments suggest a way to think about sentential relations and the effect of sentence probes. When sentences are well integrated, as they were in the present experiments, there is no superiority in noun prompt effectiveness owing to either surface or deep syntactic category, where the categories are restricted to subject and object. However, neither is there any superiority owing to the semantic role of agent and recipient. Rather, the subject of a sentence is a better prompt if it's an agent, and the object is better if it's a recipient.

The fact that there were no differences in either free sentence recall or in free noun recall suggests that subsequent differences in prompted recall are related to the structural relations to which the retrieval cues belong, rather than to the general availability of the stored meanings. The retrieval power of the nouns appears then to be associated with the structural relations they enter into, and, in particular, is most powerful when they are associated with their normal semantic and syntactic functions.

It is tempting, of course, to suggest that these results are significant for questions concerning the linguistic nature of deep structure. While the question of whether relational information contained in sentences is represented in a form more similar to case grammar relations or more similar to Chomskyan deep syntax can be raised, no unequivocal answer can be given.

Consider the two examples: (a) The Mayor publicly denounced the Governor, and (b) The Mayor withstood great pressure from the Governor. These sentences can be formally described by phrase structure grammars with NP in the first position, or as some configuration of cases (Fillmore, 1968), or as some other verb first structure (McCawley, 1970; see also Perfetti, 1972). Disregarding modifiers, which would involve embedded relations under either representation, the general form (omitting labels) of the phrase structure relations for the two sentences is: (a) ([mayor] [denounce governor]), and (b) ([mayor] [withstand pressure of governor]). A general form of the semantic relations would include: (a) denounce (AG: mayor, REC: governor), and (b) pressure (AG: governor, REC: mayor), where AG and REC are labels of the semantic relations agent and recipient. (In the case of (b), the complete semantic representation would include an embedding proposition corresponding to the assertion that the mayor did not yield to the pressure.) Thus, there are two prominent classes of hypotheses:

The Case Hypothesis assumes agent is a universal relation of significance for cognitive organization and predicts that it should provide the most powerful retrieval cue. Mayor should be best in (a), while Governor should be best in (b). The pattern of results (Table 1) expected under this hypothesis is N1 > N2 for Normal, N2 > N1 for Marked.

The Deep Syntax Hypothesis assumes the cognitive significance of the functional grammatical relations that result from the base rules of transformational grammar. It predicts that deep subject should be a more powerful prompt; thus, Mayor should be better than Governor in both (a) and (b). Thus, the pattern expected would be N1 > N2 for both Normal and Marked.

As can be seen from Table 1, the data do not support one of these hypotheses over the other for any measure. They do not even support the prediction that the two hypotheses make in common, viz. that Mayor should be a more effective prompt than Governor in sentence (a). The fact is that neither case role nor grammatical function can predict this pattern of results independently. Thus, neither linguistic hypothesis is adequate without additional assumptions. A satisfactory model of the structural relations used in language processing may have to represent information--information that is usually sharply distinguished as semantic versus syntactic--as available in some interactive (non-independent) form.

References

Anderson, J. M. The grammar of case: Towards a localistic theory. Cambridge: Cambridge University Press, 1971.

Blumenthal, A. L. Prompted recall of sentences. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 203-206.

Blumenthal, A. L., & Boakes, R. Prompted recall of sentences. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 674-675.

Fillmore, C. The case for case. In E. Back and R. T. Harms (Eds.), Universals in linguistic theory. New York: Holt, Rinehart & Winston, 1968.

Horowitz, L. M., & Prytulak, L. S. Redintegrative memory. Psychological Review, 1969, 76, 519-531.

McCawley, J. D. English as a VSO language. Language, 1970, 46, 286-299.

Perfetti, C. A. Lexical density and phrase structure depth in sentence retention. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 719-724.

Perfetti, C. A. Psychosemantics: Some cognitive aspects of structural meaning. Psychological Bulletin, 1972, 78, 241-259.

Appendix

List of Normal (N) and Marked (M) Test Sentences with Context

1. An artist was showing his paintings at a gallery.
   A critic wrote a review of the show.
   (N) The artist laughed at the opinions of the critic.
   (M) The artist suffered insults from the pen of the critic.

2. The defendant was testifying on his own behalf.
   The prosecutor wanted to ask very personal questions.
   (N) The defendant reluctantly told his story to the prosecutor.
   (M) The defendant experienced embarrassment from the questions of the prosecutor.

3. The millionaire was known for his philanthropy.
   The foundation was in great need of money.
   (N) The foundation asked for a large gift from the millionaire.
   (M) The foundation received a large gift from the millionaire.

4. The champ was in a very hard fight.
   The challenger was strong and quick.
   (N) The champ delivered a knockout punch to the challenger.
   (M) The champ sustained a head injury at the hands of the challenger.

5. A cat was prowling about the garden.
   A dog was protecting his territory.
   (N) The cat climbed a tree to escape the dog.
   (M) The cat felt terror at the barking of the dog.

6. The soldiers entered a small village.

   A farmer in the village was under suspicion.

   (N) The farmer cleverly escaped capture by the soldiers.

   (M) The farmer endured severe torture at the hands of the soldiers.

7. The old movies had a standard formula.

   A hero was introduced.

   A villain was introduced.

   (N) The hero almost always overcame the villain.

   (M) The hero went through an ordeal at the hands of the villain.

8. A customer was interested in buying a car.

   A salesman was trying to sell the most costly model.

   (N) The salesman sold an expensive car to the customer.

   (M) The salesman encountered resistance from the customer.

9. A little boy had a toothache.

   A dentist decided to remove the tooth.

   (N) The boy ran away from the dentist.

   (M) The boy felt pain at the hands of the dentist.

10. The governor was the most powerful political figure in the state.

    The mayor was known to be independent.

    (N) The mayor publicly denounced the policies of the governor.

    (M) The mayor withstood great pressure from the governor.